

UNIVERSIDADE FEDERAL DE PERNAMBUCO
CURSO ENGENHARIA DA COMPUTAÇÃO
ORIENTADOR CARLOS ALEXANDRE BARROS DE MELLO

LEANDRO HENRIQUE ESPÍNDOLA VIANA DE ALMEIDA

**Melhoria da qualidade de imagens de documentos com background
complexo**

RECIFE

2012

Sumário

1. Apresentação	3
2. Objetivos	4
3. Cronograma	4
4. Possíveis Avaliadores	4
5. Referências	5
6. Assinaturas	6

1. APRESENTAÇÃO

Documentos históricos possuem um valor significativo para a humanidade. Não só pelo fato de carregarem consigo um pouco do que realmente aconteceu no nosso passado, mas também para servir de medida sobre o quanto a sociedade moderna evoluiu, quando comparada com seus primórdios. Não raro, porém, esses documentos sofrem com a ação degradante do tempo, muitas vezes pondo em risco até mesmo o próprio conteúdo em questão.

Faz-se necessário, então, desenvolver técnicas, não só para preservar estes documentos, como também possibilitar o acesso a seu conteúdo de forma prática e eficiente, sem comprometer as informações ali contidas. Uma forma de se trabalhar um documento histórico é a digitalização do mesmo e posterior submissão às diversas metodologias existentes na área de processamento de imagens digitais.

Em se tratando de imagem digital, segundo [1], o processamento tem por objetivo melhorar a qualidade visual do material escrito para um leitor qualquer ou mesmo possibilitar uma melhor condição de percepção, reconhecimento ou armazenamento, no caso de computadores digitais.

Um passo importante para alcançar tal meta é *binarizar* a imagem, realçando o texto e desprezando as alterações de fundo e ruídos existentes no original. Uma *binarização* bem feita permite o acesso ao conteúdo do texto de forma clara e fiel, com incidência mínima de ruídos, além de facilitar o armazenamento da imagem, já que ela possui bem menos informações de cor do que a original. E permite também uma ação mais precisa dos sistemas *OCR – Optical Character Recognition* – aumentando os índices de acerto no reconhecimento de caracteres.

Nesse contexto, este trabalho tem por objetivo investigar algumas das recentes técnicas de *binarização* de imagens de documentos históricos, em particular as descritas em [2], [3] e [4], realizando experimentos sobre elas afim de se obter uma boa base de resultados para futuras melhorias.

2. OBJETIVOS

- Investigar algumas técnicas recentes de binarização de documentos históricos.
- Implementar uma ou mais técnicas, em especial [2] e [4], para realização de experimentos e estudos de caso.

3. CRONOGRAMA

A tabela abaixo ilustra o cronograma que será utilizado ao longo deste trabalho.

(A) – Revisão da Literatura

(B) – Implementação das técnicas

(C) – Experimentos e comparações

(D) – Elaboração da monografia e apresentação

	Março	Abril	Maio	Junho
(A)				
(B)				
(C)				
(D)				

4. POSSÍVEIS AVALIADORES

Tsang Ing Ren

George Darmiton

Cleber Zanchettin

5. REFERÊNCIAS

[1] GONZALEZ, Rafael C ; WOODS, Richard E. Processamento Digital de Imagens (3ª edição). São Paulo, Pearson Prentice Hall, 2010.

[2] SU, Bolan ; LU, Shijan ; TAN, Chew Lim. Binarization of Historical Document Images Using the Local Maximum and Minimum. Workshop on Document Analysis and Systems, Boston, EUA, 2010.

[3] SU, Bolan ; LU, Shijan ; TAN, Chew Lim. Document Image Binarization Using Background Estimation and Stroke Edges. International Journal on Document Analysis and Recognition, vol 13, pp. 303-314, 2010.

[4] GOVINDARAJU, Venu; SHI, Zhixin. Historical Handwritten Document Image Segmentation Using Background Light Intensity Normalization. XII SPIE Document Recognition and Retrieval, San Jose, EUA, 2005.

6. ASSINATURAS

Carlos Alexandre Barros Mello

Orientador

Leandro Henrique Espíndola Viana de Almeida

Aluno